

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-168174

(43)Date of publication of application : 14.06.1994

(51)Int.Cl.

G06F 12/00

(21)Application number : 04-319559

(71)Applicant : NEC CORP

(22)Date of filing : 30.11.1992

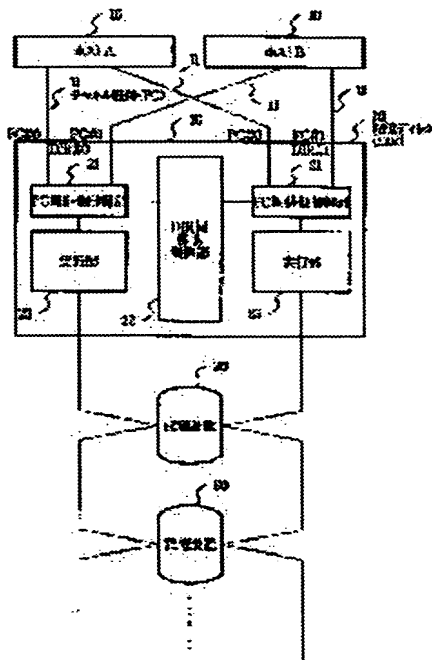
(72)Inventor : UCHIDA MITSUJIROU

## (54) FILE SUB-SYSTEM CONNECTED WITH PLURAL HOST COMPUTERS

## (57)Abstract:

PURPOSE: To prevent the stoppage of a program on first and second hosts even when the collision of an input and output command to a file device connected with the plural hosts occurs by executing the input and output command from the second host after the end of the usage of the file device by the first host.

CONSTITUTION: Each host A and B 10 executes the input and output command to a certain storage device 30 by one of a pair of channel paths (PC) 11, and the simultaneous execution by both the pair of PC11 is not operated. Then, plural memory areas provided for each storage device 30 at each of a pair of inter-PC exclusive control parts 21 are constituted of an inter-PC exclusive using memory area in which data indicating whether or not the corresponding storage device 30 is being used, and which PC 11 transmits the input and output command to the storage device 30 are stored, and a PC waiting memory area in which the data indicating that the input and output command is being waited when the corresponding storage device 30 is being used by the other PC 11 at the time of receiving the input and output command.



## LEGAL STATUS

[Date of request for examination] 30.11.1992

[Date of sending the examiner's decision of rejection] 30.01.1996

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's]

decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(Translation of selected lines)

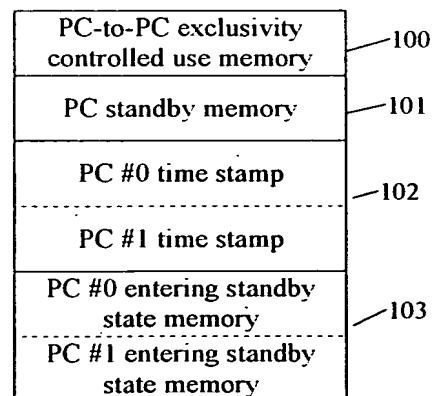
1. Japanese Patent Application, Publication No.H11-175402  
(No translation provided)

2. Japanese Patent Application, Publication No.H06-168174

[0013] Each of the above-described storage directors (hereafter called DIRs) 20 of a pair is equipped with a PC-to-PC exclusivity control unit 21 and a DIR-to-DIR exclusivity control unit 22 is provided between this pair of PC-to-PC exclusivity control units 21. The drawing provided in Fig.2 indicates a number of memory areas of a set out of a plurality of the same configuration sets, all of which are provided in either of these PC-to-PC exclusivity control units 21 and are related with the storage units 30 respectively. And one of these memory areas of the set is a PC-to-PC exclusivity-controlled-use memory area 100 and another of these memory areas of the set is a PC standby memory areas 101, wherein the former is to store data indicating whether a related storage unit 30 is under use and if so from which PC a relevant input/output instruction has been received while the latter is for storing data indicating that it is necessary for now to suspend execution of a process associated with a received input/output instruction when and if a relevant storage unit 30 is under use used by some other PC 11 when the input/output instruction is received. The drawing provided in Fig.3 indicates memory areas of a set out of a plurality of the same configuration sets, all of which are provided in the DIR-to-DIR exclusivity control unit 22 and are related with the storage units 30 respectively and the memory areas of the set comprises a DIR-to-DIR exclusivity-controlled-use memory area 110 that is for storing data for indicating whether the related storage unit 30 is under use and if so then which of the DIRs is using the storage unit 30 and a DIR standby memory area 111 that is for memorizing that a DIR holds a received input /output instruction of which the execution is suspended. Data indicating why the related storage unit 30 has come to the currently-in-use state is stored also into the PC-to-PC exclusivity-controlled-use memory area 100 and the DIR-to-DIR exclusivity-controlled-use memory area 110 when and if the related storage unit 30 is in a currently-in-use state. This implies that these two memory areas 100 and 110 store data that distinguishes the cause of the related storage unit 30 having come to be used, namely, it is because an input/output instruction is received, it is because an operation (reservation operation) for holding the storage unit 30 in the currently-in-use state is enacted after completing an input/output process, or it is because a status that is generated by a file subsystem is retained and a protection operation is in progress for protecting this status. It is possible to store at the same time standby statuses of a plurality of PCs

11 into the PC standby memory area 101 indicated in Fig.2 and standby statuses of the DIRs 20 into the DIR standby memory area 111 indicated in Fig.3. We will now explain, in the following, the operations of this particular embodiment using the flowchart presented in Fig.4 and concerned with operations of the magnetic disk sub-system associated with this embodiment in which the operations commensurate when receiving an input/output instruction from a PC #p 11.

Fig. 2



(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平6-168174

(43)公開日 平成6年(1994)6月14日

(51)Int.Cl.<sup>5</sup>

G 0 6 F 12/00

識別記号

5 3 5 Z 8526-5B

庁内整理番号

F I

技術表示箇所

審査請求 有 請求項の数 3(全 8 頁)

(21)出願番号 特願平4-319559

(22)出願日 平成4年(1992)11月30日

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 内田 密次郎

東京都港区芝五丁目7番1号日本電気株式会社内

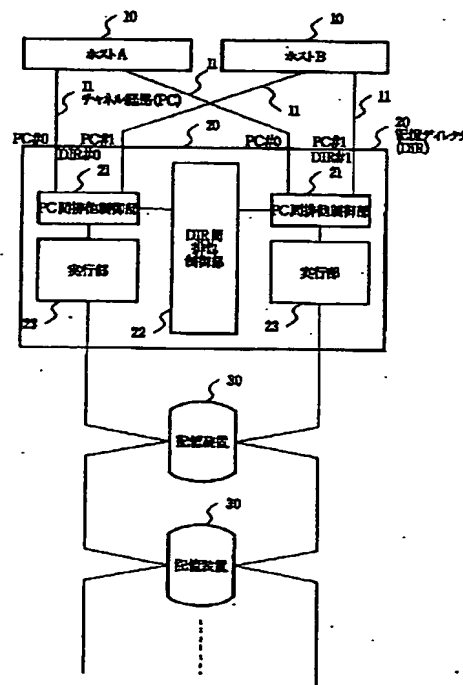
(74)代理人 弁理士 京本 直樹 (外2名)

(54)【発明の名称】 複数のホストコンピュータに接続されるファイルサブシステム

(57)【要約】

【目的】 疎結合された複数のホストに共用接続されたファイル装置への入出力指令が衝突した際に計算機システム上のプログラムが停止することを防止する。

【構成】 記憶装置30がホストA10による使用時にホストB10から入出力指令を受けた場合に、ホストB10からの入出力指令をPC間排他制御部21及びDIR間排他制御部22間にキューイングし、ホストA10による記憶装置30の使用が終了した後に、キューイングしておいたホストB10からの入出力指令を実行する。



## 【特許請求の範囲】

【請求項1】 第一および第二のホストコンピュータを含む複数のホストコンピュータと疎結合されそれら複数のホストコンピュータからの入出力指令を実行する制御装置と、前記制御装置による制御の下に前記複数のホストコンピュータにより共用されるように前記制御装置に接続されたファイル装置とを備えるファイルサブシステムにおいて、

前記ファイル装置が前記第一のホストコンピュータにより使用されている時に前記第二のホストコンピュータから入出力指令を受けた場合に、前記第二のホストコンピュータからの入出力指令を前記第一のホストコンピュータによる使用の終了まで一時的に保管（キューイング）する手段と、

前記第一のホストコンピュータにより使用中のファイル装置の使用の終了にตอบสนองして前記キューイング手段に保管された前記第二のホストコンピュータからの入出力指令を実行する手段とを備えることを特徴とするファイルサブシステム。

【請求項2】 前記キューイング手段による前記一時的保管の期間が固有のしきい値またはホストコンピュータから予め与えられたしきい値を越えたとき前記保管を解除する手段をさらに備えることを特徴とする請求項1記載のファイルサブシステム。

【請求項3】 前記キューイング手段による前記一時的保管の開始時点において前記使用中のファイル装置の使用に至った経路および使用中に至った原因を表わす情報を抽出し記憶する手段と前記一時的保管の期間が前記しきい値を越えた時点において前記ファイル装置を使用中の経路および使用中に至った原因を示す情報を抽出して記憶する手段とこれら情報抽出記憶手段の出力を前記一時的保管の解除を表わす情報に加える手段とをさらに備えることを特徴とする請求項2記載のファイルサブシステム。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、ファイルサブシステムに関し、特に疎結合された複数のホストコンピュータ（ホスト）により共用されるファイル装置を含むファイルサブシステムにおける排他使用権制御方式に関する。

## 【0002】

【従来の技術】 二つのホストAおよびBと疎結合されこれらホストAおよびBからの入出力指令を実行する制御装置と、この制御装置による制御の下にホストAおよびBにより共用されるようにこの制御装置に接続されたファイル装置とを備えるファイルサブシステムにおいて、ホストAからの入出力指令がホストBからの入出力指令とファイル装置上で非同期的に競合する。従来のファイルサブシステムでは、例えば、ホストAがファイル装置を使用中にホストBからファイル装置に入出力指令

が出されると、ホストBにファイル装置が使用中であることを報告（Busy報告）し、ホストAによる使用が終了すると、ホストBにファイル装置が使用可能に遷移したことを報告（Busy to Free報告）している。一方 Busy報告により拒否された入出力指令はホストBに保留され、Busy to Free報告を待つて再度入出力指令がホストBから発信される。上記のBusy報告、Busy to Free報告に関しては、米IBM社発行の資料GA22-6974-08, File No. S360/S370-19 "IBM System/360, and System/370 I/O Interface, Channel to Control Unit, Original Equipment, Manufacturers' Information"に記載されているのでここでは、詳述しない。

## 【0003】

【発明が解決しようとする課題】 上述した従来のファイルサブシステムは、複数の経路でホストと接続されており、また、各経路上では複数のファイル装置への入出力指令が同時に活動し得る構成を採っている。このため、あるファイル装置が使用中であるためにBusy報告を行った場合は、この使用中のファイル装置が使用可能状態に復したことを他のファイル装置に関するサービスを行っていたがために、検知できないとか、検知できても Busy to Free 報告を行えないなどの状況が生ずる。また、Busy to Free報告が行えた場合でも、そのBusy to Free報告を受け取ったホストが入出力指令を再発信した時点ではファイル装置が他のホストにより使用されている場合がある。上述のとおり、従来のファイルサブシステムでは、ファイル装置への入出力指令の発行頻度が高い場合またはファイルサブシステムの負荷が高い場合に、入出力指令のスケジューリングが破綻して実行不能な状態に陥りその入出力指令を発行したホスト上でのデータプロセッシングが停止してしまう場合がある。例えば、この入出力指令の発行元がシステム全体を制御するOSであれば、システムストールが起こってしまう。

【0004】 一方、複数の入出力指令に基づく一連のファイル変更処理が完了するまで他のホストからの入出力指令を抑止する必要があることから、入出力指令内で“入出力指令終了後もファイル装置の使用状態を継続する”操作、すなわちリザーブ操作が行われることがある。また、入出力指令にตอบสนองしてファイルサブシステムが生成したステータスの報告が何らかの原因で阻害され、そのステータスがファイルサブシステム内に保留されている間、保留中のステータスを保護するためにファイル装置の使用状態を継続する場合がある。これらの状態が何らかの原因で長時間に及ぶと、上述のデータプロセッシング停止やシステムストールが発生するが、このような現象からその原因を特定することは非常に困難である。

【0005】また、OSが入出力指令の実質的な所要時間を監視しているケースでは、長時間にわたる使用中状態が生じると、OS側でタイムオーバを検出するが、この場合の障害被疑範囲は装置障害レベルより一般的に広く、そのためシステム全体に与える影響が大きいという問題がある。

【0006】

【課題を解決するための手段】本発明の目的は、疎結合された複数のホストに共用接続されたファイル装置への入出力指令に衝突が生じてもそれらホスト上のプログラムに停止が生じないようにし、それによって上記の問題点を解消したファイルサブシステムを提供することにある。

【0007】本発明のファイルサブシステムは、複数のホストと疎結合されそれらホストからの入出力指令を実行する制御装置と、この制御装置による制御の下にそれら複数のホストにより共用されるようにこの制御装置に接続されたファイル装置とを備えるファイルサブシステムにおいて、ファイル装置が第一のホストにより使用されているあいだに第二のホストからの入出力指令を受けた場合は、第二のホストからの入出力指令をキューイングする手段を備え、第一のホストによるファイル装置の使用の終了後に前記キューイングされた第二のホストからの入出力指令を実行することを特徴とする。

【0008】また、本発明のファイルサブシステムは、上記キューイングの期間がそのサブシステム固有のしきい値またはホストが予め与えたしきい値を越えた場合に、前記キューイングされた第二のホストからの入出力指令のキューイングを解除する手段をさらに備えることを特徴とする。

【0009】さらに、本発明のファイルサブシステムは、上記キューイングの開始時点でファイル装置を使用中の経路および使用中の原因を示す情報を抽出する手段と、上記キューイングの開始からの期間が上記しきい値を越えた時点でファイル装置を使用中の経路および使用中の原因を示す情報を抽出する手段と、これら情報抽出手段の出力を上記キューイング解除ステータス情報に加える手段とをさらに備えることを特徴とする。

【0010】

【実施例】次に、図面を参照して本発明の実施例を説明する。

【0011】図1は磁気ディスク記憶装置による本発明の一実施例のファイルサブシステムの構成図、図2は図1のファイルサブシステムに用いられる制御情報を示す構成図、図3は図1のファイルサブシステムに用いられるもう一つの制御情報の構成図である。図1において、本実施例の磁気ディスク記憶装置によるファイルサブシステムは、一対の記憶ディレクタ（以降D I Rと略す）20とこれらD I R20にそれぞれ接続された複数の外部記憶装置30とで構成する。ホストA/B10の各々は、2本

のチャネル経路11（以下P C）を持ち、各P C11はD I R#020および#120にそれぞれ接続される。ホストA/B10の各々は、ある記憶装置30への入出力指令を一対のP C11のいずれか一方で実行し、それら一対のP C11の両方で同時に実行することはない。

【0012】ホストA10とホストB10は非同期的に動作しており、従ってホストA10からの入出力指令を実行中の記憶装置30にホストB10からの入出力指令が発信されることがある。

【0013】上記一対のD I R20の各々はP C間排他制御部21を備え、これら制御部21の間にはD I R間排他制御部22を配置する。図2を参照すると、上述の一対のP C間排他制御部21の各々が記憶装置30の各々について備える複数のメモリエリアが示してある。すなわち、それらメモリエリアは、対応記憶装置30が使用中であるか否かおよび使用中である場合はP C11のどれからの入出力指令により使用中であるかを表示するデータを記憶するP C間排他使用メモリエリア100 と、入出力指令を受信した時点で対応記憶装置30が他のP C11で使用中であった場合にこの入出力指令の実行を待機中であることを表示するデータを記憶するP C待機メモリエリア101 とから成る。図3を参照すると、D I R間排他制御部22が記憶装置30の各々について備えているメモリエリアは、対応記憶装置30が使用中であるか否かおよび使用中である場合にいずれのD I Rが使用中であることを表示するデータを記憶するD I R間排他使用メモリエリア110 と、D I Rが待機中の入出力指令を保有していることを記憶するD I R待機メモリエリア111 とから成る。上述のP C間排他使用メモリエリア100 およびD I R間排他使用メモリエリア110 には、記憶装置30が使用中である場合にその使用中の状態に至った原因を表示するデータを併せて記憶する。すなわち、入出力指令を受けたために現時点で実行中であるのか、入出力動作の終了後に記憶装置30の使用状態を継続する操作（リザーブ）が実行されたために使用中であるのか、あるいは入出力指令を受けたのちファイルサブシステムが生成したステータスが保留されていてこのステータスを保護する操作が実行されたために使用中であるのかの区別を明らかに表示するデータをこれらメモリエリア110 および111 に記憶する。尚、図2に示したP C待機メモリエリア101 および図3に示したD I R待機メモリエリア111 には、複数のP C11およびD I R20の待機状況をそれぞれ同時に記憶できる。次に、本実施例の磁気ディスクサブシステムのP C#p11から入出力指令を受信した時のD I R#d20の動作のフローチャートを示す図4を参照して本実施例の動作を説明する。

【0014】D I R#d20は、P C間排他使用メモリエリア100 およびD I R間排他使用メモリエリア110 を順次調査し（ステップ210, 211）、いずれかの排他使用メモリエリア100, 110 が“使用中”であれば、D I R#d

20のPC待機メモリエリア101に“PC#p待機中”、DIR待機メモリエリア111に“DIR#d待機中”を登録して(ステップ213)、入出力指令の実行を待機する。両方の排他使用メモリエリア100,110が“空き”であれば、PC間排他使用メモリエリア101を“PC#p使用中”、DIR間排他使用メモリエリア111を“DIR#d使用中”として(ステップ212)、入出力指令を実行する。

【0015】上述の磁気ディスクサブシステムのPC#p11からの入出力指令の実行の終了時のDIR#d20の動作を示すフローチャートである図5を参照すると、

(DIR#d'は他のDIR20を、PC#p'は他のPC11を意味する)、DIR#d20は、DIR待機メモリエリア110を調査する(ステップ221,222)。DIR待機メモリエリア110が“DIR#d'待機中”であれば、DIR排他使用メモリエリア110を“DIR#d'使用中”とし、PC間排他使用メモリエリア100を“空き”とする(ステップ224)。DIR待機メモリエリア111に“DIR#d(自DIR)待機中”が登録されていれば、DIR排他使用メモリエリア110の内容を維持してPC間排他使用メモリエリア100を“PC#p'使用中”とする(ステップ225)。DIR待機メモリエリア111が“待機無し”であれば、PC間排他使用メモリエリア100とDIR間排他使用メモリエリア110の両方を“空き”とする(ステップ223)。上記磁気ディスクサブシステムの入出力指令の実行を待機しているDIR#d20の動作のフローチャートである図6を参照すると、DIR#d20はまずPC待機メモリエリア101を調査する(ステップ231)。待機中のPC11が無ければ、無動作である。待機中のPC11があれば、PC間排他使用メモリエリア100を判定する(ステップ232)。PC間排他使用メモリエリア100が“空き”であれば、さらにDIR間排他使用メモリエリア110を調査する(ステップ233)。DIR間排他使用メモリエリア110が“DIR#d使用中”であれば、PC間排他使用メモリエリア100を待機中のPC#11で使用中とし(ステップ236)、PC待機メモリエリア101およびDIR待機メモリエリア111を解除して(ステップ237)、待機中の入出力指令を実行する。またはPC間排他使用メモリエリア100上の使用中PC#11とPC待機メモリエリア101上の待機中PC#11が一致していれば、PC待機メモリエリア101およびDIR待機メモリエリア111を解除して(ステップ237)、待機中の入出力指令を実行する。複数のホストからの入出力指令に競合があった場合の本実施例におけるキューイング制御は上述のとおり行われる。尚、DIR20は20ms間隔でカウントアップするタイマカウンタを備える。従って、入出力指令にすぐ応答できなくてその指令を待機状態にした時、タイマカウンタ値をPC間排他制御部21のPC#pタイムスタンプ102に記憶し(図4、ステップ214)、またこの時点にお

ける使用中の経路および使用中に至った原因を示す情報をPC間排他制御部22のPC#p待機開始メモリエリア103に記憶する(図4、ステップ215)。上記の入出力指令を待機中にしている間に、現在のタイムスタンプ値とPC#pタイムスタンプ102値との差分により待機期間を算出し、待機期間がサブシステムの固有のしきい値またはホストから予め与えられたしきい値を越えた場合は待機中の入出力指令を異常終了させる。ホスト10への異常終了報告ステータス内に、待機開始時点でのPC#p待機開始メモリエリア103の内容およびしきい値オーバー発生時点の使用中の経路と使用中の原因を示す情報を取り込むこともできる。本実施例は、2個のDIR20と2個のPC11とを備えているが、これら構成要素が3個以上であって待機中の入出力指令が複数ある場合は、ラウンドロビンあるいは待機行列等の手法で待機中の2以上のPC11からひとつを選択してこれに使用権を与えることは容易である。

【0016】尚、本実施例においては、DIR間排他制御部を一对のDIRの間に配置した例を示したが、この制御部の構成要素を磁気ディスク記憶装置の各々に分散配置することも可能である。

【0017】

【発明の効果】以上説明したとおり、本発明のファイルサブシステムは、複数のホストからの入出力指令を順次に行うことができるので、各ファイル装置への入出力指令の発信頻度が高い場合でもそれら入出力指令を確実に実行できる。従って、計算機システム上のプロセスを停止させることなく円滑に実行できる効果を奏する。

【0018】また、疎結合された複数のホストにより共用されるファイル装置の“使用中状態”の時間的長さを監視する機構を備えているので、その“使用中状態”の発生原因の如何に関わらず入出力指令動作を継続し、停止させない。また、障害が発生した場合は、最悪でもその障害をファイル装置障害レベルに留め、システム全体への影響を防ぐことができる。さらに、この監視機構を設けたことにより、きめ細かい情報の採取が可能となり、“使用中状態”が長時間にわたっている経路の情報、“使用中状態”に至った原因の種別の情報などを入出力指令の異常終了ステータスに含め、これにより異常発生時の原因の究明が容易になる。

【図面の簡単な説明】

【図1】本発明の一実施例である磁気ディスクサブシステムの構成図である。

【図2】図1のファイルサブシステムに用いられる制御情報の構成図である。

【図3】図1のファイルサブシステムに用いられるもう一つの制御情報の構成図である。

【図4】本実施例のPC#p11から入出力指令を受信した時のDIR#d20の動作のフローチャートである。

【図5】本実施例のPC#p11から発信された入出力指

令の実行が終了した時のDIR#d20の動作のフローチャートである。

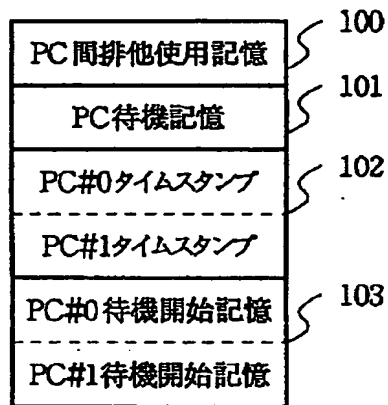
【図6】本実施例の入出力指令の実行を待機しているDIR#d20の動作のフローチャートである。

【符号の説明】

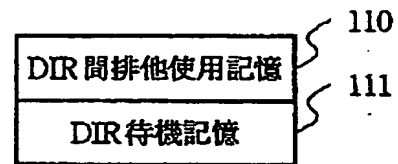
- 10 ホスト
- 11 チャンネル経路 (PC)
- 20 記憶ディレクタ (DIR)
- 21 PC間排他制御部

- 22 DIR間排他制御部
- 23 実行部
- 30 記憶装置
- 100 PC間排他使用メモリエリア
- 101 PC待機メモリエリア
- 102 PC#pタイムスタンプ
- 103 PC#p待機開始メモリエリア
- 110 DIR間排他使用メモリエリア
- 111 DIR待機メモリエリア

【図2】

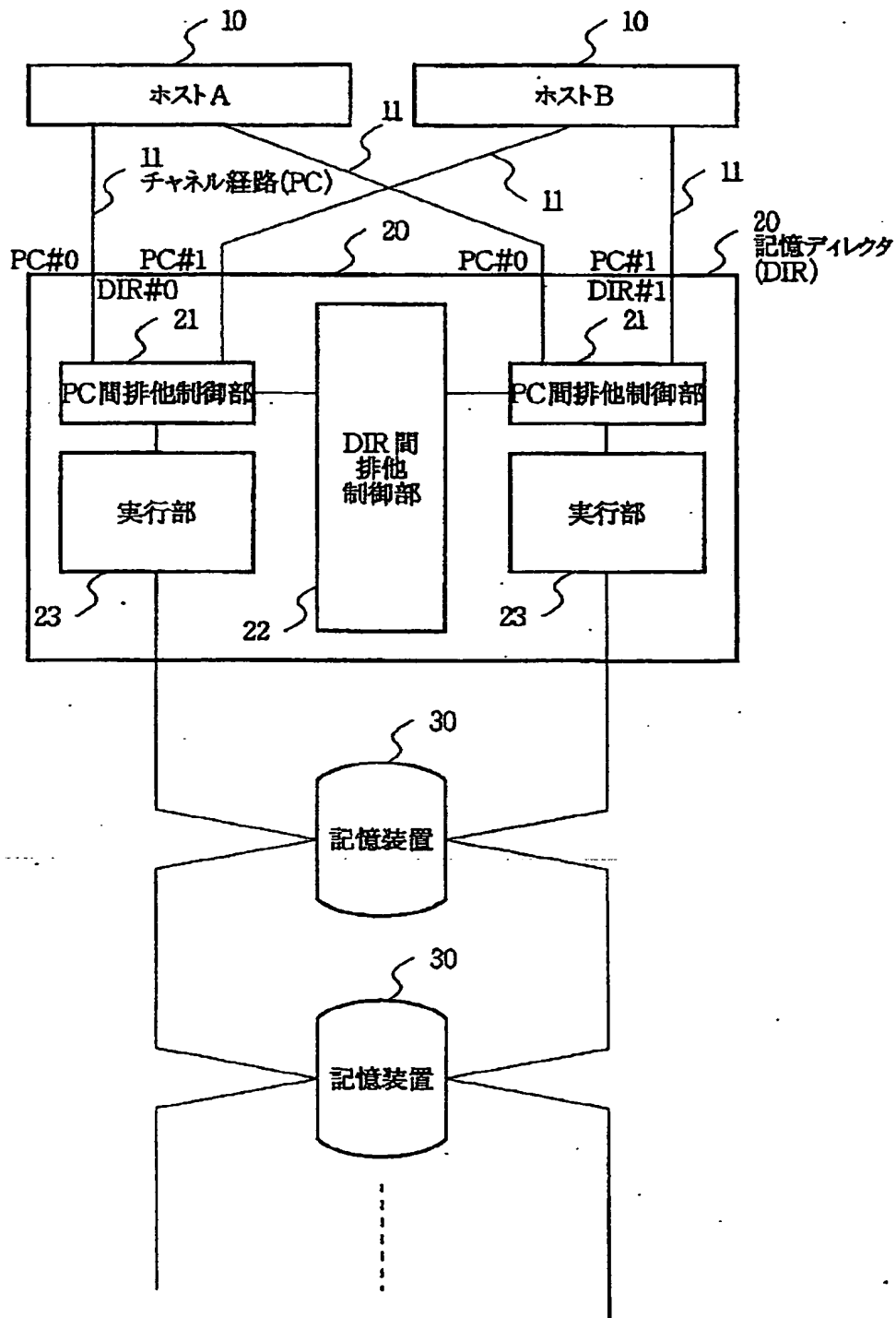


【図3】

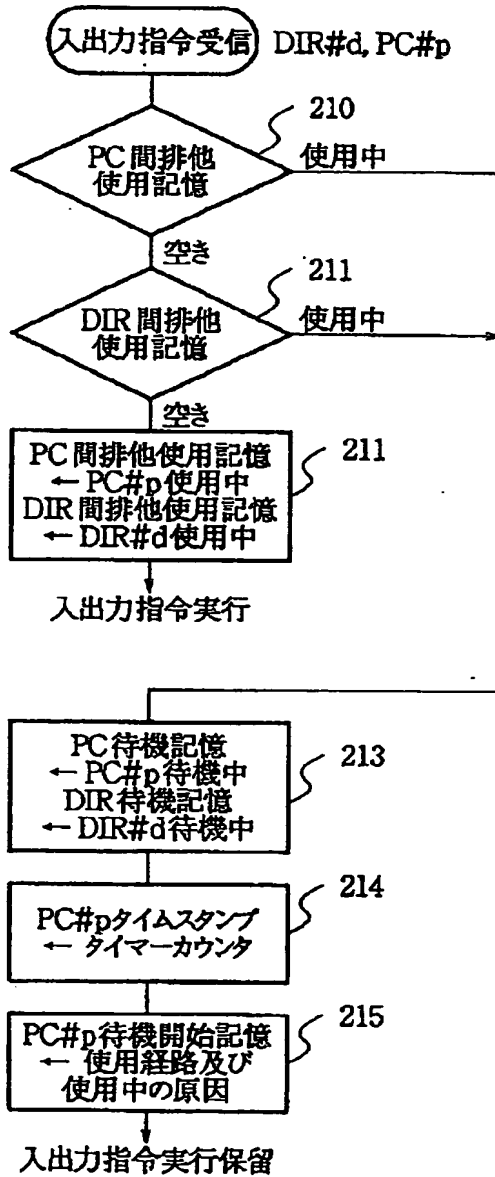




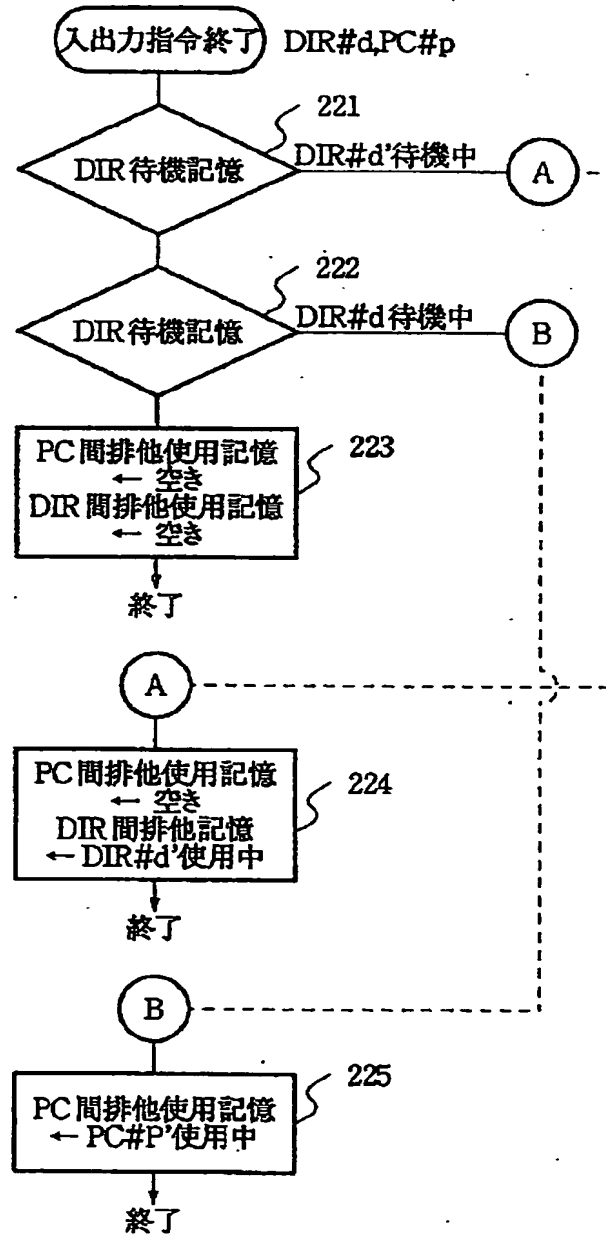
【図1】



【図 4】



【図 5】



【図 6】

